# White Paper Report

Report ID: 102718

Application Number: PR5012211

Project Director: Gregory Wilsbacher (gregw@sc.edu)

Institution: University of South Carolina Research Foundation

Reporting Period: 4/1/2011-9/30/2013

Report Due: 12/31/2013

Date Submitted: 12/18/2013

# AEO-Light (Optical Sound Extraction Software): a White Paper

by:

Greg Wilsbacher

Submitted to the National Endowment for the Humanities
Preservation Access Research and Development

December 18, 2013

## INTRODUCTION

Sound motion picture film combines two distinct information streams. Intermittent systems record and project the image information. In camera, the film is advanced into the gate, stopped for a short time, exposed, and then advanced again. Projection requires a similar movement: each frame is stopped in the gate (typically for 1/24th of a second) as a shutter rotates to release the light and throw the image onto a screen. In contrast, sound is recorded as a continuous stream of information. In the projection environment the sound information (whether optical, magnetic or digital) is read by a sound head over which the film is driven at a continuous pace.

Bringing these two distinct information systems together successfully took decades of experimentation. Once the basic technical conventions were established, a stable division of labor was also produced. The craft of creating finished audio was distinct from the craft of shooting and editing a finished image sequence, and a third type of expertise was necessary to combine the two. This industrial process has left a difficult legacy for film preservationists, who must reproduce it (if separate audio and image source elements exist) or reverse-engineer it (if a finished sound-film survives). Sound increases not only the technical complexity, but also the expense, of film preservation.

Within the last five years, the development of increasingly affordable frame-integrity film scanners capable of imaging film at resolutions approximating the quality of photochemical reproduction has provided new tools for image preservation. Although long-term archival storage of large digital cinema files remains a major hurdle, the digital intermediate (DI) process allows for the creation of a durable photochemical film element without the generational loss of information or the cost of intermediate film elements.

With the traditional industrial process in mind, engineers initially designed film scanners to digitize only the image content within the frame. Images would be captured intermittently, and sound could be handled separately. In any case, to achieve full resolution required scanning rates well below the 24 fps necessary to read the audio track in a traditional fashion. By 2009, some scanner designs (like the Kinetta archival film scanner) allowed digitization of the film strip from one edge to another. Inclusion of the the optical sound track in the recorded image opened the door to a new approach to sound film preservation.

The AEO-Light project was founded on the idea that audio and image information both could be extracted from the same digital scan of an optical motion picture sound film. This would decrease the expense of preservation and reduce archives' reliance on commercial laboratories to perform basic preservation tasks. As a result, more of the twentieth century's rich film legacy would escape the dustbin of history.

The resulting software breaks new ground in film preservation. It demonstrates that audio information can be properly extracted from image files created through an intermittent process. In addition to promising savings of time and labor, the process eliminates the requirement of scanning at 24 frames per second. Optical sound can be recovered from film scans made at any speed. Although scanners continue to require slower speeds to achieve the highest resolution, it has become possible to imagine that improved sensors, transport mechanisms, and software may soon allow digitization to exceed the "speed of sound". Finally, the AEO-Light software enables audio to be extracted from film scans at any time—dissociating the audio extraction work from the act of image digitization. Scanned film including the optical sound track can have audio extracted from it years after the scan was made. For the first time, uncompressed .tif files can be viewed as a preservation file for both image and sound, creating, in the process, a bulwark against migrating more complex media files from one format to another. The open source nature of the AEO-Light project serves this longer vision by sharing its source code and encouraging community development and stewardship of these technologies so that they are available to preservationists in the future.

## PROJECT SCOPE

The AEO-Light project is a collaboration between the University of South Carolina's Moving Image Research Collections (MIRC) and its Interdisciplinary Mathematics Institute (IMI). It was made possible by funding from the National Endowment for the Humanities' Preservation and Access Research and Development award program with additional support from the University's College of Arts and Sciences.[1] Professional consulting services were provided by Bob Heiber of Chace Audio by Deluxe with additional in-kind services provided by Chace Audio by Deluxe and Film Technology, Inc.

The project has three broad goals. Foremost, it seeks to digitize and preserve sound information from optical tracks imaged during the film scanning process. In addition, though, it seeks to encourage the creation of digital surrogates (full edge-to-edge scans) of physical film objects to increase the possibility that future preservationists and scholars can have access to all the information content of motion picture film stock in the absence of the physical object. Finally, the project seeks to promote the development of essential preservation tools as open source software. Obviously, the project team only has control over the first of these goals, but it is hoped that the success of AEO-Light software will impact the others.

AEO-Light is designed to be independent of the scanning process so that its functionality can be preserved indefinitely even as scanning systems evolve. This scanner neutral approach means supporting the ingest of many video and individual frame formats. AEO-Light is also designed to export synchronized digital video as well as a free-standing audio file and to provide

[1] Members of the grant team were: Greg Wilsbacher (MIRC, Principle Investigator), Borislav Karaivanov (IMI, Co-Principle Investigator), Pencho Petrushev (IMI, Co-PI), Mark Cooper (MIRC, Co-PI), L. Scott Johnson (IMI, programer) and Brittany Braddock (MIRC, scanning technician). This white paper draws upon the work of the entire team. Special thanks to Ashely Blewer for providing pro bono graphic design services and to others whose input improved the quality of AEO-Light (Shai Drori, Robert Gilmore, Inna Kozlov, Jeff Kreines, Diana Little, Amilcar do Carmo Lucas, Jeff Mather, Alexander Petukhov, and Paisa Seeluangsawat).

fundamental preservation metadata about the exported files to facilitate ingest into asset management systems.

AEO-Light can extract audio from most optical sound tracks from any gauge of film. In general, optical tracks fall into one of two broad types: variable area and variable density. The former encodes electric sound onto film stock by varying the intensity of the light in response to the frequency of the signal.[2] The latter encodes the electric sound by varying the area of emulsion exposed (see Figures 1 and 2).

Figure 1: 35mm, variable density sound

Figure 2: 16mm, variable area sound

AEO-Light is also able to process simple 2-channel stereo tracks. Additional development will be necessary in order to properly extract audio from intermediate noise reduction tracks and Dolby stereo tracks.

---

[2] Electric sound is sound converted from its mechanical form into an electric signal by way of a condenser microphone and then subsequently processed by an amplifier prior to being sent to the optical sound system.

FUNCTIONAL REQUIREMENTS

AEO-Light extracts audio from film scans that meet the following requirements:

- The scans must be made so that the film imaged includes the optical soundtrack in addition to the image-frame.
- The scans must also be configured so that some information above and below each image-frame is included. Figure 3 shows a sample frame with a generous amount of vertical overlap. In practice, a smaller amount of overlap is adequate for audio extraction.

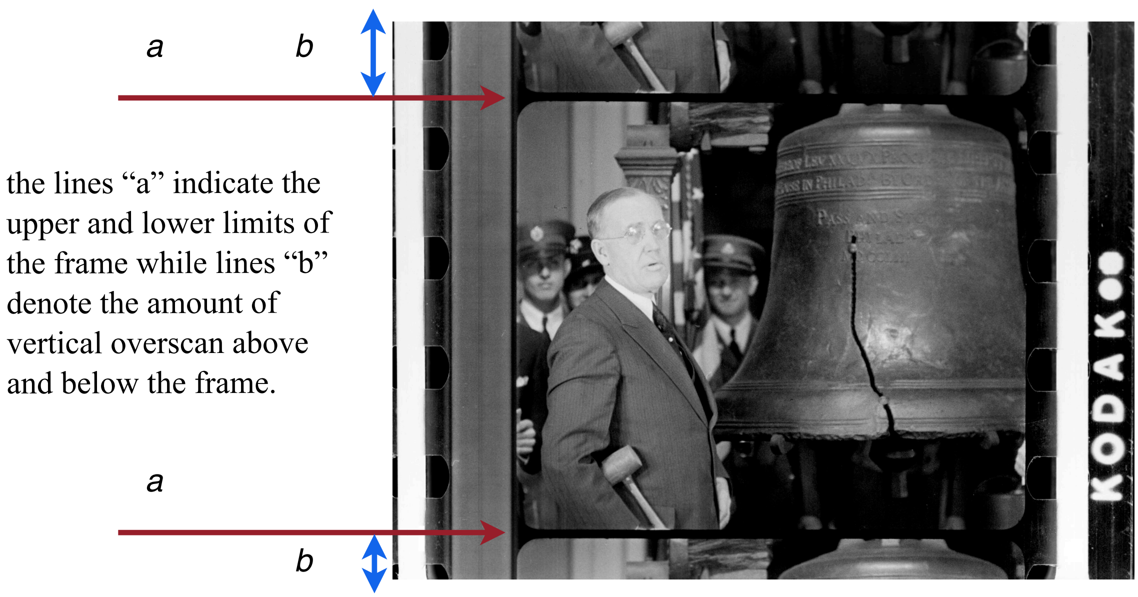


Figure 3: Sample frame illustrating the nature of the overscan required by AEO-Light.

- The scans must contain enough resolution to provide meaningful audio information. The minimum resolution required to produce acceptable audio is as yet undetermined. Users are encouraged to scan at the highest resolutions possible for initial tests.
- AEO-Light was not originally designed to process sound-only tracks, but tests indicate such tracks can successfully be extracted.

Hardware and software requirements for AEO-Light:

- 64-bit Windows, Mac OS/X, Linux
- AEO-Light application, http://sourceforge.net/projects/aeolight/?source=directory

• Matlab Compiler Runtime (MCR) 2013a (v.8.1) or better, http://www.mathworks.com/products/compiler/mcr/
• FFmpeg v. 0.11 or later (required for video export functionality). See **http://ffmpeg.org** for documentation and downloads. AEO-Light beta has been tested against the static builds provided by Tessus (Mac) at http://www.evermeet.cx/ffmpeg/ and Zeranoe (Win) at http://ffmpeg.zeranoe.com/builds/ Users unfamiliar with FFmpeg are encouraged to install one of the Windows or Mac static builds.

In practice AEO-Light does not make substantial demands on CPU processing power. However, AEO-Light does provide for parallel processing on multiple CPU cores. Performance is also improved as more physical RAM is made available. AEO-Light is a read/write intensive workflow in its current design. Because solid state drives reduce read/write times they provide better performance over systems utilizing traditional spinning disk drives.

## EXTRACTION METHODOLOGY

AEO-Light is the result of two years of research and experimentation aimed at developing a tool able to extract audio from files created by practically any model of scanner provided the scanner can meet the modest horizontal and vertical overscan requirements outlined above. Being in effect *scanner neutral* means accommodating the potential anomalies of each scanner not just as designed and delivered by the manufacturer but also as they develop over time with use. AEO-Light has routines that provide for two condition variables that might negatively impact the resulting audio: (1) imperfections in illumination in the gate area and (2) unsteadiness in the film scan itself. Imperfections in illumination are addressed through optional calibration routines prior to the creation of audio signals. Unsteadiness in a scan increases the complexity of correctly registering the frame line between each frame pair in the film. AEO-Light provides three options for registering frame lines to improve registration in the presence of unsteady scans.

The complete extraction methodology is as follows.

1) Read frame information from source material;[3]
2) Define region of frame with audio information;
3) Extract sound signal from each frame;
4) Generate and apply a calibration mask (optional);
5) Compute guess for most likely frame pair overlap (optional);
6) Search for the best-matching overlap for each frame pair;
7) Combine the individual sound signals into a complete audio signal.

[3] Here and elsewhere in this document, frame refers to the over-scanned image that includes what is typically referred to as a film frame as well as the extra information from the optical sound track area and the small portions of the preceding and following frames.

*Step One*: AEO-Light can process discrete frame files (e.g., .dpx or .tif) directly or it can process frame information fed to it by the MATLAB video engine or by FFmpeg. It first attempts to read the information in MATLAB, turning to FFmpeg in instances when MATLAB fails. AEO-Light's capacity to process video is, thus, limited only by the A/V libraries available to MATLAB and FFmpeg.

*Step Two*: Because the optical sound track exists within defined parameters on the film, its left and right bounds can be established with relative ease. The movement of the film throughout the scanning process necessitates, though, a sampling routine to ensure that the left and right bounds established are applicable for the entire sequences of frames to be processed. The degree of freedom allowable for the establishment of the soundtrack bounds is related to the general type of optical track being processed. Variable density tracks, for example, allow for very narrow bounding box. In contrast, variable area tracks restrict the width of the bounding box to that required to capture the full modulation of the sound wave.

The user is presented a series of frames selected at random with vertical lines superimposed on them. For each frame, the user drags the lines to the edges of the soundtrack area to indicate the bounds of the soundtrack. After the user provides bounding box parameters for a sample of frames, the intersection of the selected regions is formed and its bounds are assumed acceptable for soundtracks areas on all frames.

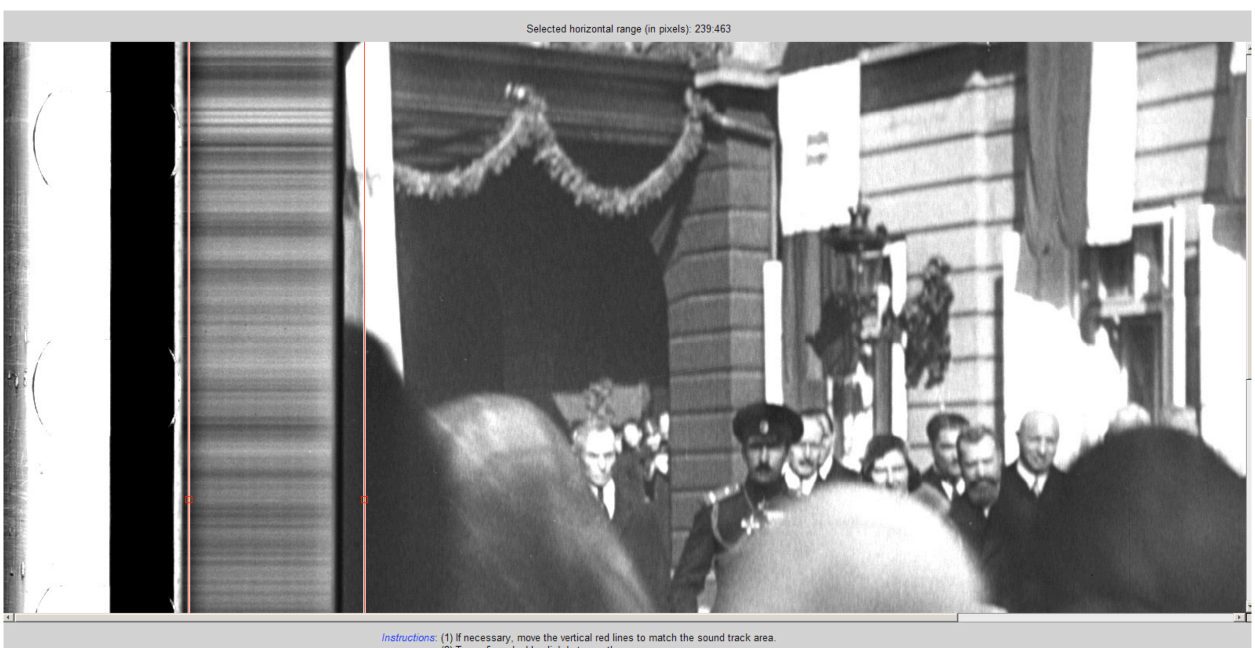


Figure 4: Soundtrack bounds are determined interactively by moving draggable vertical lines (in red) superimposed on the image.

There is no limit to the number of bounding boxes that can be set by the user. Simple stereo (two channel) tracks can be extracted using this process by setting two bounding boxes in such a way as to isolate one channel from another. In the case of variable density tracks that manifest physical damage, multiple bounding boxes can provide iterations of the audio track sampled from different regions of the overall track area. These iterations can then provide audio restorationists with richer source material from which a final audio track can be drawn.

*Step Three*: A sound signal from each frame is extracted. Once the soundtrack is located, the process is straightforward and applies without further refinement to most types of tracks. At this stage, the software has created a sound signal for each frame of the film selected for processing.

*Step Four*: This optional step creates a calibration mask to compensate for the signal distortion caused by imperfect illumination in the gate. After experimenting with a number of calibration methods a *sound-based sound calibration* method was integrated into AEO-Light.

*Sound-based sound calibration* recovers the pattern of light by looking at the average distortion of the individual audio signals and then creating a mask to correct that distortion. The individual audio signals extracted from images are averaged over the whole series to give a single short signal, a sound-based average, which is smoother than the individual signals and captures the pattern of their mutual relation. In an ideal scan, the average of a large number of sound signals would be a more or less flat signal. Therefore, in a sound-based sound calibration any deviation from being flat must be due entirely to unevenness in illuminating light and not inherent to the sound signal encoded optically on the film. Logic dictates that the patten of signal distortion in the sound signals is a good approximation of the aberrations in the lamphouse illumination.

Once the mask is created it is improved by either smoothing via a *moving average* routine or by a *polynomial fit* routine. Smoothing the sound-based average should bring the average closer to the actuality of the illumination pattern since it can be expected that the illumination varies subtly in space (see Figure 5). An additive 1D sound-based calibration mask is produced by negating and properly shifting the smoothed average. Each of the extracted sound signals is corrected by applying the calibration mask (see Figure 6). The calibrated signals have a flat average (or close to it depending on the quality of the mask produced), mirroring the signal as it would have been produced in uniform illumination.

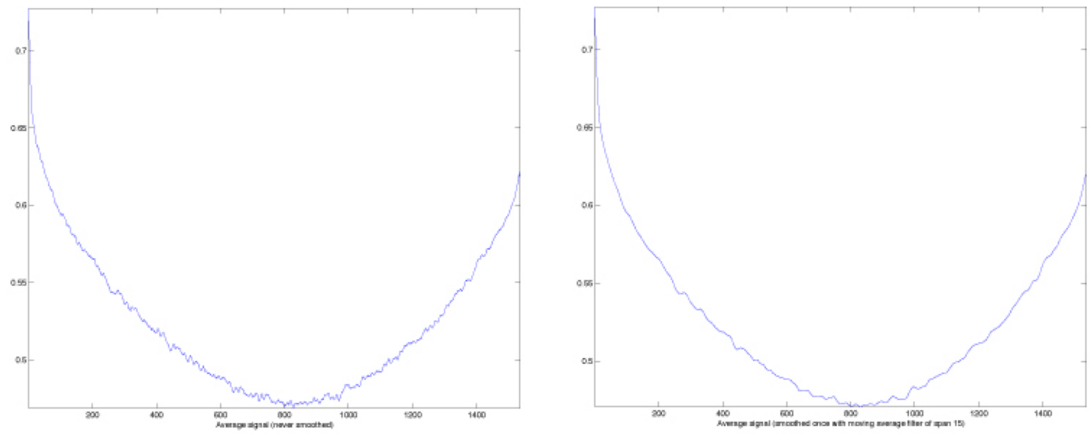


Figure 5: Unsmoothed (left) and smoothed (right) signal calibration mask.

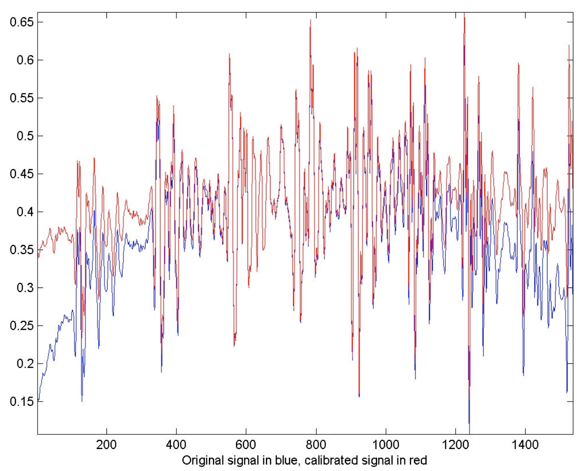


Figure 6: Original sound signal in blue; calibrated signal in red.

*Step Five*: This optional step aims at reducing the number of calculations required to register an accurate frame overlap for each consecutive pair of frames (step 6). In an ideal scan the frame overlap would be consistent from one frame pair to another. However, the reality of scanning introduces small to significant variations in frame overlap from one frame pair to another: these may derive from flaws in the film stock itself to deficiencies in the scanner system. The amounts of overlap are expected to vary within a small range (20 pixels, for example) with the understanding that some variations well outside this range must be accommodated.

This step operates on a small number of pairs of consecutive sound signals (200, for example) drawn at random. All possible overlaps between the two signals in a pair are considered. These range from overlapping by a single pixel to completely overlapping each other. For each overlap a relative difference is computed as the norm of difference between the two overlapping portions divided by the sum of their norms. Different norms can be employed in these computations but the city-block norm appears as the least expensive among those quantifying the signals' relation in a stable way. Each overlap is ranked according to the number of pairs having a local minimum of the relative difference at the overlap. The highest ranking overlap is our guess for the one most likely to occur.

*Step Six*: Precise calculation of the frame overlaps is essential for acceptable audio. Even small misalignments can create audible anomalies or (worse) distort the timing of the extracted audio signal so that it doesn't match the length of the original optical sound track. Three methods for computing the best overlap are here considered: exhaustive search, local search, and Fast Fourier Transform (FFT). Each of the three is applicable to images and, when properly restricted, to sound signals as well.

Every two consecutive images share a common area of film. In order to align them the second image has to be shifted in relationship to the first image. This shift is quantified by a two-dimensional vector specifying the amount of horizontal and vertical translation. It is assumed that the natural position corresponding to the zero shift vector is where the bottom edge of the first image coincides exactly with the top edge of the second. In this setting, the horizontal component of a shift vector could assume both negative and positive integer values depending on whether the second image is shifted to the left or to the right of the first image. The vertical component of a shift vector ranges from one (representing a single pixel row of overlap) to the image height (representing complete overlap).

The exhaustive search takes into consideration all admissible shifts of the second image with respect of the first. For each shift vector, a relative difference is computed as the norm of difference between the two overlapping portions divided by the sum of their norms. The norm could be the city-block norm or any other norm that is computationally feasible and is stable even when presented with pixel outliers (e.g., noise or other artifacts introduced by scanning or compression or similar). The shift vector for which the smallest relative difference is attained gives the best-matching overlap. Reformulating the method for sound signals is straightforward: the horizontal component of the shift vectors is discarded, leaving just the vertical shift, and relative differences are computed using sound signals instead of images. Clearly, the exhaustive search method requires substantial computation.

The local search method reduces the computational cost of the exhaustive search. It uses an initial guess for the shift vector that has a zero horizontal component and a vertical component equal to the guess for the most likely (vertical) overlap computed in step 5. Then the same kind of relative differences are computed, but locally, in a small neighborhood around the guessed shift vector. Calculations are carried out only for those shift vectors whose city-block distance to the guessed shift vector is not greater than a fixed radius (10 pixels, for example). If the shift vector for which the smallest relative difference is attained is strictly inside of the described neighborhood, then it is considered the the best match. If the shift vector is outside the described neighborhood, then the global minimum is attained on the boundary of the analyzed neighborhood, and a reliable conclusion cannot be drawn without further investigation. In this case, the size of the search neighborhood is increased and the process is repeated until the global minimum no longer occurs on the boundary.  Reformulation for sound is completely analogous to that of the exhaustive search.

The Fast Fourier Transform (FFT) provides an alternative means of identifying the frame overlaps by analyzing the sound signals extracted from each frame rather than any image information. The Fourier transform of the convolution of two consecutive images is calculated as the pointwise product of their Fourier transforms. These transforms are computed very efficiently using the Fast Fourier Transform algorithm.

*Step Seven*: The last step is the most straightforward of all. Once the correct amount of overlap for each pair of consecutive sound signals is determined, the two pieces are merged. Within the overlapping portion the combined signal is formed as a weighted average of the two signals. Different weights can be used, but as a rule they need to decay smoothly as samples move away from the middle of the signal. This property ensures a smooth transition from one sound signal to the next while giving more credibility to the sound samples extracted from the central portions of the images, and less to those based on pixels near image boundaries (see Figure 7)

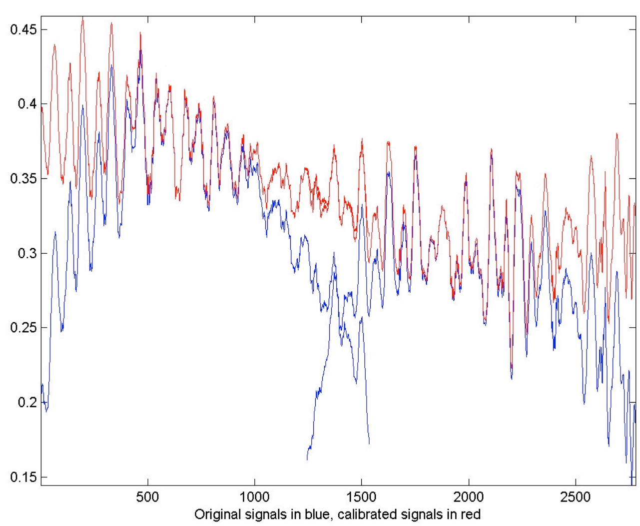


Figure 7:  An overlapping frame pair, showing both the uncalibrated and the calibrated signals

## REMAINING CHALLENGES

At the time of its initial public release, AEO-Light provides quality audio from scans of motion picture film meeting the basic criteria for an acceptable scan of motion picture film.[4] The software performs best when these scans are of positive audio tracks, especially those of release prints, and early user feedback suggests that AEO-Light produces high quality audio more reliably from 16mm films.

Future development of AEO-Light will expand on its current capabilities by:

* Increasing the quality of audio overall with special attention to 35mm film.
* Improving performance characteristics with the aim of faster than real time audio extraction.
* Implementing new algorithms to address the specific characteristics of optical sound tracks (negative variable density, Dolby stereo, etc.).
* Ensuring the viability of the source code over time by migrating to an open development environment.

### *Increasing the quality of audio*

Because AEO-Light assembles intermittent audio information pulled from each frame into a virtual band of continuous information *prior* to converting that information into digital audio, the precision of the overlap for each frame pair weighs heavily on audio quality. Improving the frame pair overlap routines to increase the precision of these calculations generally will enable AEO-Light audio to provide consistently preservation quality audio characterized by a low signal to noise ratio as well as complete sampling of the frequency range.[5]

Currently, AEO-Light performs well with scans that are very steady and contain sound that varies in frequency, pitch, complexity, etc. (as is the case with most films). In these cases the Fast Fourier Transform routine, which determines overlap only by analyzing the audio signal, can effectively locate the overlap for each frame pair because the signal has complex characteristics that make validating the frame line position easier. However, the software's performance is known to degrade in the presence of steady tones, silence, or other instances in which the audio signal displays a consistent, unvarying pattern.

Test films produced to assess the performance of audio systems are essential to tracking and improving AEO-Light's performance in this area. A sweep tone test film was used to map

---

[4] See Functional Requirements.

[5] The quality of the original optical track plays a significant role in determining what can count for preservation quality audio. The aim of AEO-Light is not to provide tools for sound restoration but to provide a sound signal of sufficient quality to allow for restoration if desired.

frequency response and noise characteristics.[6] Chace Audio by Deluxe processed the test film through its proprietary optical to digital conversion system and provided the resulting audio file as a quality benchmark.

Spectrographic analysis provides a visual reference of one full sweep (100 Hz to 10kHz) from the test film. Figure 8 shows the acoustic energy of the sweep as represented in the benchmark audio measured in hertz (y-axis) against time (x-axis). The band of acoustic energy steps gradually through the frequency range of the sweep with nicely defined subharmonics. What noise that is present exists in the sub-audible, lower frequencies.

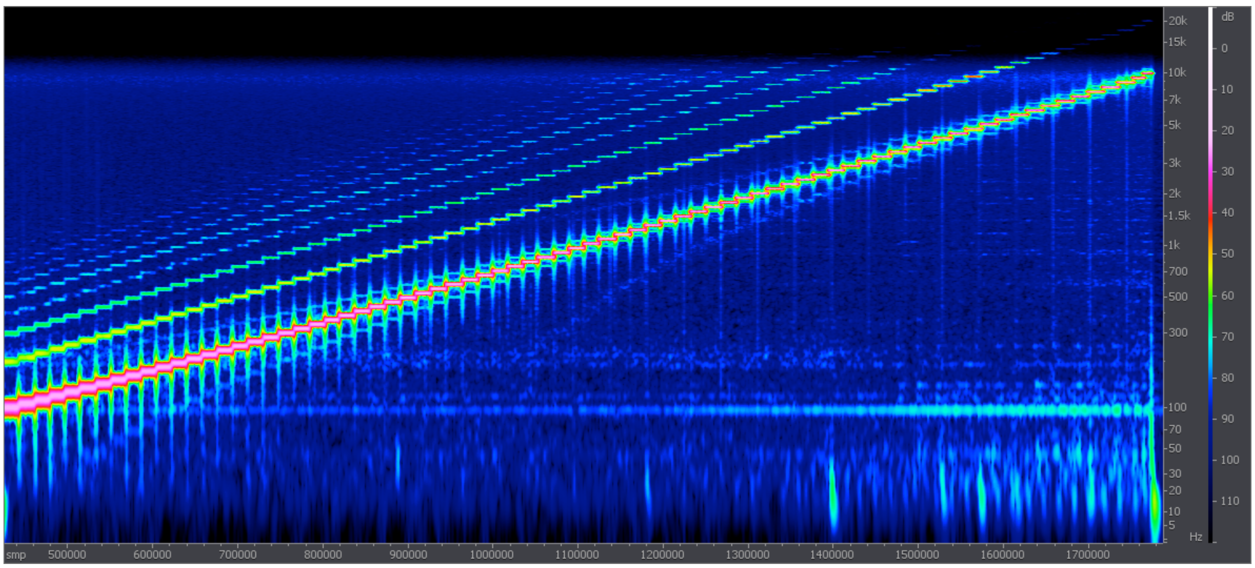


Figure 8: Spectrograph of Chace Audio by Deluxe sweep tone

Figure 9 is a representation of a sweep from a scan of the same test film element processed by AEO-Light. The AEO-Light audio compares unfavorably in two ways. First, the lower frequencies show considerable noise. While much of this falls below the 20 Hz audible threshold, there is a significant band of acoustic energy at 100 Hz with lighter bands up to approximately 250 Hz. More importantly, though, the AEO-Light spectrograph reveals an excessive amount of noise moving through the frequency range of the sweep. This noise is most pronounced between 500 Hz and 3000 Hz.

[6] Thanks go to Chace Audio by Deluxe and Film Technology, Inc. for providing this and other test films to the project.

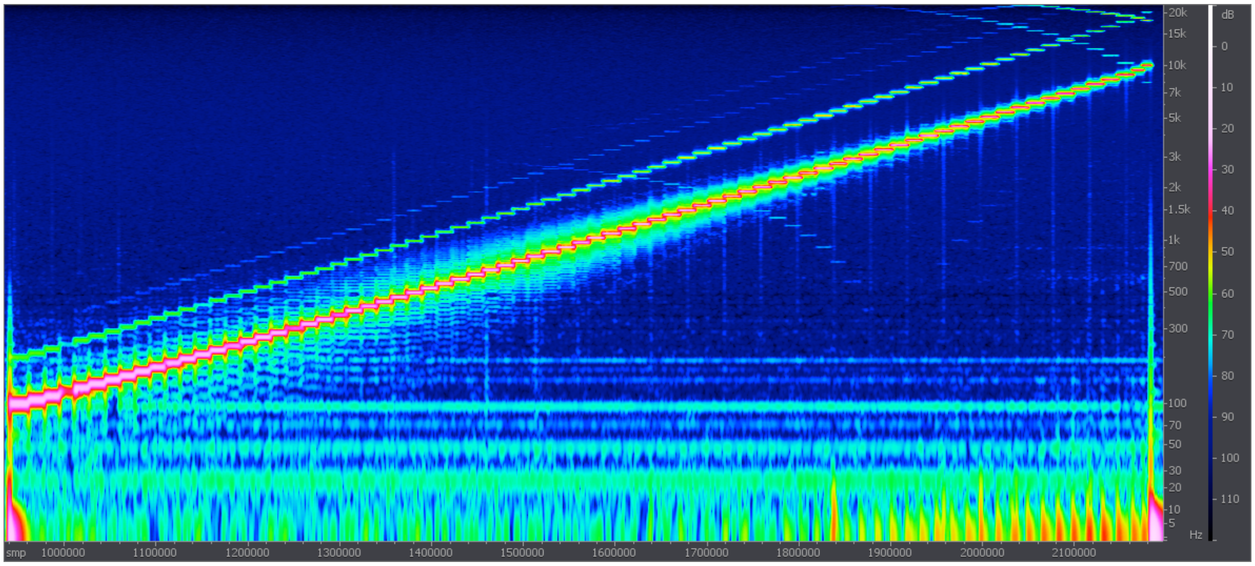


Figure 9: Spectrograph of AEO-Light produced sweep tone.

In addition to the undesirable noise, Figure 9 shows that the harmonics of the sweep do not match those of the Chace benchmark. There are seven harmonics of decreasing energy in the Chace extraction, but only three in the AEO-Light audio. Figure 9 also shows an audible descending sub-harmonic (2500 to 700 Hz).

Although the sweep tone presents variation in sound, the individual steps of increasing frequency are instances of unvarying sound and as a result these spectrographs provide a good indication of the impact of less-than-ideal frame pair overlap in some AEO-Light audio extractions.

*Improving Performance Characteristics*

Early AEO-Light designs focused on routines that compensated for lamphouse illumination irregularities and stabilized frame sequences to allow for reliable calculation of frame pair overlaps. These routines were time-consuming (ranging from 3 seconds per frame to 1 fps) and were not considered practical. Revisions to these routines increased speeds, and the addition of a Fast Fourier Transform routine pushed the speed (for Windows 7) to about 15 fps. Although this speed is still below run time for sound film (24 fps), it is fast enough for practical use.

To have the greatest preservation impact, however, AEO-Light should have faster throughput. Bringing this performance to what is desired by the community, i.e., near instantaneous extraction of audio once all parameters are set, will require wholesale revision of the source code, moving the algorithms out of its current development environment, MATLAB, into a compiled language such as C++. Although the MATLAB code is compiled, it is still executed through a virtual machine on the end user's computer. This creates a performance bottle neck that can't be overcome by revising MATLAB routines. This bottleneck is readily apparent when the performance of Windows and Mac variants of AEO-Light are compared because the MATLAB virtual machine is optimized more for Windows operating systems than for Mac or Linux operating systems. Table 1 below provides a comparison of Mac and Windows.[7]

| **AEO-Light 0.9** | dpx (2K) | video (2K) | video (3K) |
|---|---|---|---|
| Windows 7 | 3.1 fps | 12.5 fps | 10.6 fps |
| Mac OS | 2.6 fps | 3.2 fps | 3.58 fps |

Table 1: Performance Characteristics of Windows and Mac versions of AEO-Light on comparable systems.

Recently MATLAB provided support for GPU acceleration. While this new capability has not yet been written into revised AEO-Light routines, its incorporation in the future will certainly improve performance, but to what degree can't be ascertained in advance.

*Implementing new algorithms to address the specific characteristics of optical sound tracks*

At present, AEO-Light approaches each optical track in a neutral fashion. It doesn't register whether the track is variable area or variable density, whether it is a positive or a negative image, etc. In the most general theory of extraction, the algorithms don't require such information in order to convert the pixel information contained with the region defined by the bounding box. In this, it mirrors the activity of the photo electric cell of a traditional optical sound head, which will read as sound whatever film passes over it. Unlike a projector, though, AEO-Light is expected to extract sound from films found throughout the archive, including negative and intermediate stock.

Providing for the accurate negative-to-positive conversion of variable density tracks is the most important of these additional algorithms. Conversion of negative to positive in the digital realm often assumes a linear relationship between pixel values. Such an assumption works well for variable area tracks precisely because the tracks encode information in a binary manner. The shape of the sound wave is defined by absolutes: exposed/unexposed, black/white, one/zero.

In contrast, variable density tracks are defined by the variations of exposure as opposed to the area exposed. As a result, the sound information of variable density tracks is intimately related to the characteristic curve (H & D curve) of the film stock, which is logarithmic, not linear. In a properly recorded negative density track the incoming audio signal is engineered so that the positive print from the negative reflects the desired audio signal. Variable density negatives processed by the current build of AEO-Light sound acceptable but they lack an accurate

---

[7] These speeds are taken from scans of the same motion picture film (35mm sweep tone test). Note that the 3K scan processed more quickly than the 2K scan on the Mac OS. The 3K scan was made by Media Preserve and was steadier than the 2K scan produced at the University of South Carolina.

representation of the audio signal in the upper and lower frequencies, those regions corresponding to the head and toe of the characteristic curve. The distinction between emulsion characteristics desired for variable density recording as distinct from variable area recording are clearly shown by Eastman Kodak's own sensitometry for its specialized optical sound recording stocks as manufactured in 1942, a time in which both density and area tracks were still commonly in use (see Figure 10).

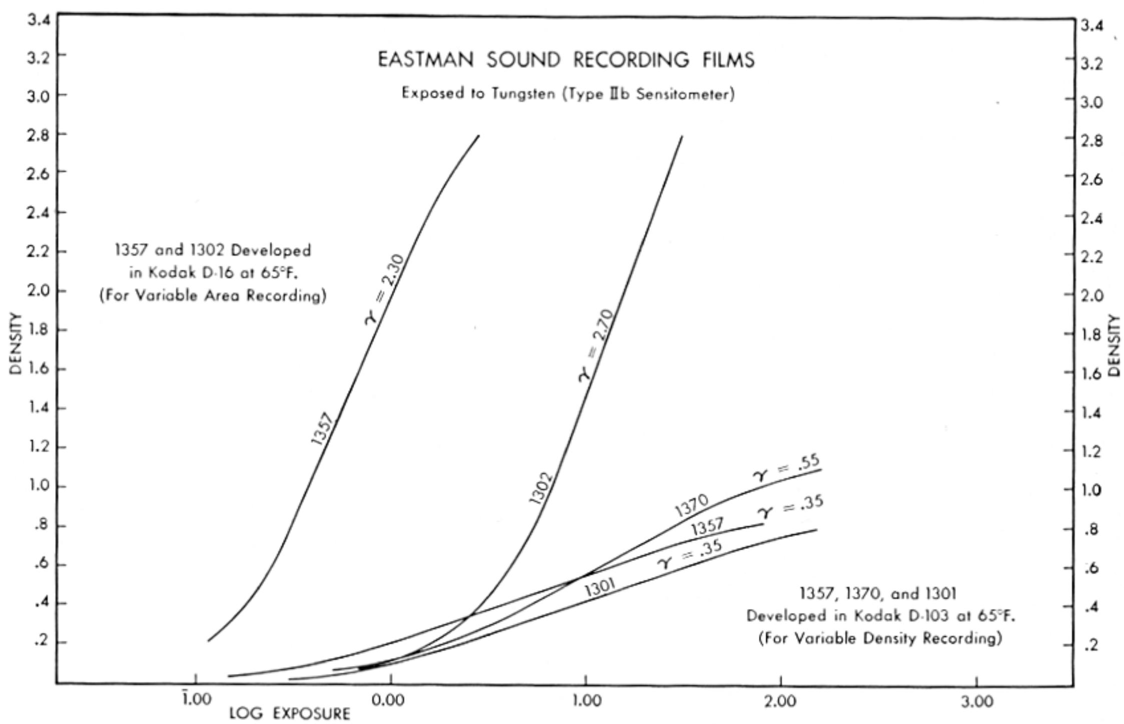


Figure 10: Sensitometry for Eastman optical sound recording films.[8]

Providing an algorithm that can emulate the photochemical conversion of negative to positive would increase the viability of AEO-Light as the film sound extraction tool of choice. Implementing such a feature, though, will require incorporating a representative number of characteristic curves into the algorithm so that users can identify the curve that best approximates that of the original film stock.

In addition to providing algorithms designed specifically for variable density negative, AEO-Light builds in the future should incorporate algorithms able to properly extract sound from early intermediate noise reduction tracks, often called "push-pull" tracks (see Figure 11). Unlike negative density tracks, push-pull tracks require an additional noise reduction process and *do not sound acceptable* when read directly by an optical sound head or by AEO-Light.

---

[8] *Eastman MotionPicture Films for Professional Use* (Rochester: Eastman Kodak Company, 1942), 64.

Figure 11: 35mm, push-pull track.

Properly extracting push-pull tracks is feasible and should require a routine that corrects the phase shifts implemented during the noise reduction process. Information about push-pull recording is readily available in the Academy's *Motion Picture Sound Engineering.*[9]

It may also be desirable to facilitate the preservation of Dolby stereo tracks by enabling AEO-Light to convert the two visible tracks of a Dolby stereo recording into the four channel *passive matrix* (left, right, center and surround).[10] While the resulting four channel sound will still lack the final polish of sound processed by the proprietary *active matrix* Dolby hardware decoder, the extraction of the four channels as discrete streams would be an important first step for the open source long-term digital preservation of Dolby stereo sound.

*Ensuring the viability of the source code*

AEO-Light was established as an open source project to provide the film preservation community with long-term access to tools necessary for the permanent preservation of our shared film history. Currently the source code as well as compiled versions for Windows, Mac and Linux reside on Sourceforge, http://sourceforge.net/projects/aeolight/?source=directory.

While the source code is shared openly, the development environment for the code is not entirely open source.[11] AEO-Light does utilize the power of open-source FFmpeg for transcoding audio and video and to provide additional support for video ingest. However, all the core algorithms

---

[9] See Fred Albin, "Noise Reduction," in *Motion Picture Sound Engineering* (New York: D. Van Nostrand Company, 1938).

[10] John Polito (of Audio Mechanics) provided a basic primer in Dolby Stereo during his 2103 "Reel Thing" talk, "Roots and Stems: Superior Practice in REmastering Stereophonic Cinema " (paper presented at the annual meeting of the Association of Moving Image Archivists, Richmond, VA, November 2013).

[11] While MATLAB is not open source, an open source project compatible with MATLAB is being developed (see, http://www.gnu.org/software/octave/).

for audio extraction as well as the GUI and other supporting routines are written in MATLAB, a proprietary (and somewhat expensive) suite of tools. MATLAB (http://www.mathworks.com/products/matlab/) was selected as the development language to leverage its powerful toolkits that streamline solutions to complex mathematical calculations. MATLAB proved its usefulness by allowing rapid implementation of experimental processes.

True open source sustainability will only come when the current MATLAB source code is rewritten into one of the development environments more compatible with the open source community.

## Summary

AEO-Light extracts audio from scans of motion picture films with optical sound tracks. In its current state, the software provides high quality reference audio tracks of most motion picture sound films. In some cases the audio extracted may be considered of high enough quality to function as a preservation audio track, especially in the case of 16mm films. The software innovates by allowing for the preservation of two distinct information systems (intermittent image and continuous sound) in one pass with edge-to-edge scanning and by capturing time-dependent audio information at any speed--24 fps is no longer the fixed speed for film sound transfer.

While the software is designed for use in the present to unpack the audio information from film scans, the software's innovative approach and open source nature creates the possibility of using uncompressed still images (e.g., .tif files) as the preservation master format for many optical sound films. Tif files from digital film surrogates will contain not only the motion picture image content, but also the optical sound information. By proving that audio can be extracted from such files, AEO-Light demonstrates that audio can be extracted from still image files any time in the future. While such capability would not replace the digital preservation and management of complex time-based media files combining image and sound, it would provide a bulwark against loss of content decades in the future as .tif files are more sustainable over time than moving image media files.

To date, the AEO-Light project has accomplished much. However, continued development of AEO-Light is important to further improve audio quality and to expand the types of optical sound that can be successfully preserved.